POPULATION FILE(S)

Areas covered: *Each SEER Registry, the 50 States, and the Total U.S.*

Years: *1969-1997*

Races: *White, Black, and Non-White (1969-1997)*
White, Black, American Indian/Aleutian Islander/Eskimo, and Asian/Pacific Islander (1990-1997)
White Non-Hispanic, White Hispanic, Non-White Non-Hispanic, and Non-White Hispanic (1990-1997)

Documentation Version: *April 2000*

OVERVIEW

The SEER population files have been constructed for the analysis of SEER and other databases. They contain population data for a specified set of years, geographic areas and demographic characteristics. The source of these populations is estimates for intracensal years for special racial groups provided by the U.S. Census Bureau and estimates computed by NCI. The title page specifies the geographic areas, racial groups, and years for which populations are provided.

This documentation defines the record format, field definitions, and file characteristics for any population file.

STORAGE CHARACTERISTICS

A. Size: dependent upon file content — for most files, may be approximated by:
           # of geographic areas * # of races * 2
           (where the 2 represents male plus female)

           (Records for specific races may not be included if when census taken, no persons of that race resided in that specified geographic area or estimates were not computed.)

B. Organization: Sequential

FILE CHARACTERISTICS

The file has the following characteristics:

A. Logical Record Length: 200 characters
B. Record format: Fixed length records
C. ASCII text file

RECORD IDENTIFIER

Record Identifier: Population Year/Month of Census/Geographic Area/Race/Sex where Geographic Area may be a combination of the SEER Registry, State, County, or Total U.S.

There are six types of population files that this documentation pertains to. Each contains slight differences in how the geographic fields (state, registry, and county) are used to identify the populations. Below describes how these fields are used with each type of file.

A.   SEER REGISTRY FILES
1.   Filename: BYREGISTRY.TXT
2   Location: SEER Public-Use CD-ROM
3.   Races: Dependent on directory containing file.
4.   The registry based records are identified by the SEER Registry codes (01-35) in the registry field, the two-character alphabetic U.S. Post Office abbreviation in the state field, and " 000" in the county field.

B.   SEER COUNTY FILES
1.   Filename: BYCOUNTY.TXT
2.   Location: SEER Public-Use CD-ROM
3.   Races: Dependent on directory containing file.
4.   The county based records are identified by the SEER Registry codes (01-35) in the registry field, and the FIPS state and county code in the state and county fields.

C.   TOTAL UNITED STATES FILE
1.   Filename: US_6996.TXT
2.   Location: SEER Web site
3.   Races: Total Black, Total White, Total Non-White
4.   The total United States records have "98" in the registry field, "US" in the state field, and "000" in the county field.

D.   INDIVIDUAL STATE FILES
1.   Filename: <AA>6996.TXT (where AA is alphabetical abbreviation of state)
2.   Location: SEER Web site
3.   Races: Total Black, Total White, Total Non-White
4.   FIPS state and county codes are used in the state and county fields. The registry field contains "98".

E.   STATE TOTALS FILE
1.   Filename: ST_6996.TXT
2.   Location: SEER Web site
3.   Races: Total Black, Total White, Total Non-White
4.   FIPS state codes are used in the state field. The registry field contains "98" and the county field contains "000".

F.   STANDARD MILLION FILE
1.   Filename: STDMIL.TXT
2.   Location: SEER Public-Use CD-ROM and SEER Web site
3.   The standard population year is the only applicable field.

The populations contained in a SEER registry record (entire state or combination of counties) and the sum of the populations of each county within a SEER area (i.e., sum of the SEER county records) are not equivalent for years prior to 1980. This is due to the way in which 1971-1979 estimates were developed for minority populations in the registry records.

ITEM NAME	ITEM #	POSITIONS	LENGTH
RECORD IDENTIFICATION			
Population Year	01	001-002	2
Race	02	003-004	2
SEER Registry	03	005-006	2
State	04	007-008	2
Standard Metropolitan Statistical Area (SMSA)	05	009-012	4
County	06	013-015	3
Census Tract	07	016-021	6
Month of Census	08	022-022	1
Sex	09	023-023	1
Filler — zero	10	024-024	1
Standard Population Year	11	025-026	2
Filler — space	12	027-027	1
Suppression Indicator	13	028-028	1
FIVE YEAR POPULATIONS			
Population: Ages <5	14	029-037	9
Population: Ages 05-09	15	038-046	9
.	.	.	.
.	.	.	.
.	.	.	.
Population: Ages 80-84	30	173-181	9
Population: Ages 85+	31	182-190	9
POPULATION: ALL AGES	32	191-200	10

NAME: Population Year

ITEM NO: 001

PURPOSE: To specify the year to which the population applies

LENGTH: 2

CONTENTS: 73 : =   1973
                74 : =   1974
                75 : =   1975
                      .
                      .
                      .
                80 : =   1980
                81 : =   1981
                      .
                      .
                      .

NAME: Race

ITEM NO: 002

PURPOSE: To specify the race/Spanish origin of the population represented

LENGTH: 2

CONTENTS: 12 : =   Total Black
                20 : =   Total White
                30 : =   Total Non-White

                04 : =   American Indian, Aleutian Islander or Eskimo
                05 : =   Asian or Pacific Islander
                12 : =   Total Black
                20 : =   Total White

                01 : =   White Non-Hispanic
                02 : =   White Hispanic
                10 : =   Non-White Non-Hispanic
                11 : =   Non-White Hispanic

*** The values that occur in this field are dependent on the context of the file. ***

NAME: SEER Registry

ITEM NO: 003

PURPOSE: To specify the SEER registry to which the population applies

LENGTH: 2

CONTENTS: 00 : =   Not Applicable
                   01 : =   San Francisco-Oakland SMSA
                   02 : =   Connecticut
                   20 : =   Metropolitan Detroit
                   21 : =   Hawaii
                   22 : =   Iowa
                   23 : =   New Mexico
                   25 : =   Seattle (Puget Sound)
                   26 : =   Utah
                   27 : =   Metropolitan Atlanta
                   31 : =   San Jose-Monterey
                   35 : =   Los Angeles
                   98 : =   United States

NAME: State

ITEM NO: 004

PURPOSE: To specify the state of the population represented

LENGTH: 2

CONTENTS: zero : = not applicable

ss : =   the two-digit Federal Information Processing Standards (FIPS) state codes:

01	Alabama	30	Montana
02	Alaska	31	Nebraska
04	Arizona	32	Nevada
05	Arkansas	33	New Hampshire
06	California	34	New Jersey
08	Colorado	35	New Mexico
09	Connecticut	36	New York
10	Delaware	37	North Carolina
11	District of	38	North Dakota
	Columbia	39	Ohio
12	Florida	40	Oklahoma
13	Georgia	41	Oregon
15	Hawaii	42	Pennsylvania
16	Idaho	44	Rhode Island
17	Illinois	45	South Carolina
18	Indiana	46	South Dakota
19	Iowa	47	Tennessee
20	Kansas	48	Texas
21	Kentucky	49	Utah
22	Louisiana	50	Vermont
23	Maine	51	Virginia
24	Maryland	53	Washington
25	Massachusetts	54	West Virginia
26	Michigan	55	Wisconsin
27	Minnesota	56	Wyoming
28	Mississippi		
29	Missouri		

st : =   the two-character alphabetic U.S. Post Office abbreviations (used with SEER registry data):

US	United States	HA	Hawaii
AR	Arizona	MI	Michigan
CA	California	NJ	New Jersey
CT	Connecticut	NM	New Mexico
GA	Georgia	UT	Utah
IA	Iowa	WA	Washington

NAME: Standard Metropolitan Statistical Area (SMSA)

ITEM NO: 005

PURPOSE: To specify the Standard Metropolitan Statistical Area
of the population represented

An SMSA is defined as a large population nucleus and nearby communities which have a
high degree of economic and social integration with that nucleus.

LENGTH: 4

CONTENTS: zero : = not applicable

smsa : = the four-digit FIPS code identifying the SMSA

See the U.S. Census Bureau document *PHC80-R5 Geographic Identification Code Scheme*
for definitions of SMSA codes.

NAME: County

ITEM NO: 006

PURPOSE: To specify the county of the population represented

A county is defined as the primary political and administrative subdivision of a state.

LENGTH: 3

CONTENTS: 999, zero : = not applicable

ccc : = the three-digit FIPS county code — counties are uniquely identified within a
state

For a complete list of the counties and their assigned codes for each state, see the
Worldwide Geographic Location Codes, published by the U.S. General Services
Administration, Finance Office, in June, 1982.

NAME: Census Tract

ITEM NO: 007

PURPOSE: To specify the census tract of the population represented

A census tract is defined as a small statistical subdivision of a county.

LENGTH: 6

CONTENTS: zero : =   not applicable

9999V99 : =   the four-digit basic tract code which may be followed by a two-digit suffix code where V is an assumed decimal point

If no suffix is present, two spaces are used. For example, a census tract consisting of only a four-digit basic code would appear as 0001bb (where b is a space). A census tract consisting of the basic code and a suffix would appear as 000101.

NAME: Month of Census

ITEM NO: 008

PURPOSE: To specify the month in which the census was taken

LENGTH: 1

CONTENTS: 4 : =   April — the decennial censuses of 1970, 1980 and 1990 were completed during April

7 : =   intracensal estimates were computed for July 1 of the year

NAME: Sex

ITEM NO: 009

PURPOSE: To specify the sex of the population represented

LENGTH: 1

CONTENTS: 1 : =   Male
                   2 : =   Female

NAME: Standard Population Year

ITEM NO: 011

PURPOSE: To specify the year represented for records containing standard million populations

LENGTH: 2

CONTENTS: space : =   not applicable
                     40   : =   1940 U.S. Standard Million
                     50   : =   1950 U.S. Standard Million
                     60   : =   1960 U.S. Standard Million
                     70   : =   1970 U.S. Standard Million
                     80   : =   1980 U.S. Standard Million
                     90   : =   1990 U.S. Standard Million
                     00   : =   2000 U.S. Standard Million (5-year groups)
                     WR : =   World Standard Million
                     CA : =   1991 Canadian Standard Million)
                     EU : =   European Standard Million (5-year groups)

        *** This field is only used in the Standard Million file. The field will be blank in the other files. ***

NAME: Suppression Indicator

ITEM NO: 013

PURPOSE: To indicate if suppression exists or existed in the populations for this race and state/county

LENGTH: 1

CONTENTS: 0 : = no suppression
                   1 : = all populations suppressed (all cells contain zero)
                   2 : = populations originally suppressed — derived from existing populations for county

NAME: Population — 5 year age group

ITEM NOS: 014-031

PURPOSE: To specify the number of persons in the specified age group residing in the geographic area of the specified race/Spanish origin and sex

LENGTH: 9 for each age group

CONTENTS: zero — 999999999 : =   population

NAME: Population: All Ages

ITEM NO: 032

PURPOSE: To specify the total number of persons residing in the geographic area of the specified race/Spanish origin and sex

LENGTH: 10

CONTENTS: zero — 9999999999 : =   population